The Geometry and Matching of Curves in Multiple Views

Cordelia Schmid, Andrew Zisserman

The geometry and matching of curves in multiple views

Cordelia Schmid[1] and Andrew Zisserman[2]

[1] INRIA Rhône-Alpes,655 av. de l'Europe, 38330 Montbonnot,France
[2] Dept of Engineering Science, 19 Parks Rd, Oxford OX1 3PJ, UK

Abstract. In this paper there are two innovations. First, the geometry of imaged curves is developed in two and three views. A set of results are given for both conics and non-algebraic curves. It is shown that the homography between the images induced by the plane of the curve can be computed from two views given only the epipolar geometry, and that the trifocal tensor can be used to transfer a conic or the curvature from two views to a third.
The second innovation is an algorithm for automatically matching individual curves between images. The algorithm uses both photometric information and the multiple view geometric results. For image pairs the homography facilitates the computation of a neighbourhood cross-correlation based matching score for putative curve correspondences. For image triplets cross-correlation matching scores are used in conjunction with curve transfer based on the trifocal geometry to disambiguate matches. Algorithms are developed for both short and wide baselines. The algorithms are robust to deficiencies in the curve segment extraction and partial occlusion.
Experimental results are given for image pairs and triplets, for varying motions between views, and for different scene types. The method is applicable to curve matching in stereo and trinocular rigs, and as a starting point for curve matching through monocular image sequences.

1 Introduction

This paper has two strands. The first is set of novel geometric results for curves in two and three views. The results are developed for both algebraic (conics) and non-algebraic curves. The key unifying idea is that the plane of the geometric entity can be computed from two views. The plane can then be used to transfer that entity into a third view via the trifocal tensor [8, 22, 23]. The transfer of points and lines via the trifocal tensor is well known and explored. This paper extends transfer to conics and curvature.

The second strand is an application of these results to the automatic matching of curves over multiple views. The idea here is that in addition to the geometric constraints that must be satisfied by matched curves, there are also photometric constraints on the intensity neighbourhood of the curve. These constraints are realised by computing intensity neighbourhood cross-correlation mediated by the homography (projective transformation) induced by the plane of the curve.

This paper can justly be seen as extending two papers by Faugeras and Roberts. The first, [7], dealt with the transfer of conics and curvature from two views to a third via epipolar transfer; in this paper the transfer is via the trifocal tensor and so does not suffer from the failings of epipolar transfer [28]. The second, [18], dealt with the trinocular matching of curves using only geometric constraints; in this paper these constraints are augmented by photometric constraints based on neighbourhood intensity correlation. In our previous work [20] this photometric constraint was shown to be a powerful disambiguation measure for line matches over two and three views.

1.1 Related literature on curve matching

Curve matching is a reoccurring topic in the stereo and motion literature, with the main application being the recovery of 3D geometry. A subsidiary goal has been to qualitatively distinguish between surface curves and apparent contours [2, 25]. Curves are a natural next step after lines, since they are a more faithful representation of a contour which would otherwise be represented by a polygonal chain. In this paper we are not considering closely spaced views where snake tracking is viable.

Basic criteria for curve matching in stereo pairs were established by Pollard *et al* in the PMF Stereo Algorithm [16]. Their primary criteria were the epipolar and ordering constraints, but these were supplemented by figural continuity to overcome problems where curves coincide with epipolar lines (since if the curve coincides with an epipolar line, point correspondences cannot be determined). Zhang and Gerbrands [27] again used epipolar geometry as their primary constraint, but followed this with criteria on the variation in disparity along the putatively matched curves. Brint and Brady [4] matched curves in trinocular views. Their primary matching constraint was a similarity measure based on the deformation between a curve and its putative match in another view. This measure was used to eliminate many potential mismatches. Trinocular consistency constraints were only used as a final verification. An alternative to matching curves individually is to first carry out a monocular grouping, and then match the groups. This is the approach adopted by Chung and Nevatia [5], and continued in work at USC by Havaldar and Medioni [10]. In this paper we match curves individually without first grouping.

Nayar and Bolle [15] developed a photometric matching constraint based on the ratio of intensities across region boundaries. This ratio of intensities has good invariance to lighting conditions. However, it is less general than the neighbourhood intensity correlation presented in this paper because it only applies to homogeneous regions, and it is also less discriminating.

1.2 Overview

Section 2 describes the geometry of conics and general curves in two and three views. Conic and curvature transfer methods are given, and these methods are assessed in section 3 by an implementation on real images. Matching is described

in section 4. Algorithms are given for two and three views, and for short and wide baselines.

Two sequences are used to evaluate the geometry and matching. These are: a "plate" sequence, figure 2; and a "bottle" sequence, figure 3. Examples on other sequences are given in [21]. Frames are selected from these sequences to form image pairs or triplets. The camera motion between the frames is fairly uniform, so that the frame number is a good indicator of the distance between views.

The matching method developed in this paper requires only projective multiple view relations: the fundamental matrix for two views and the trifocal tensor for three. Camera calibration is not required. These relations are computed automatically from point (corner) correspondences [1, 24, 26].

2 Geometry

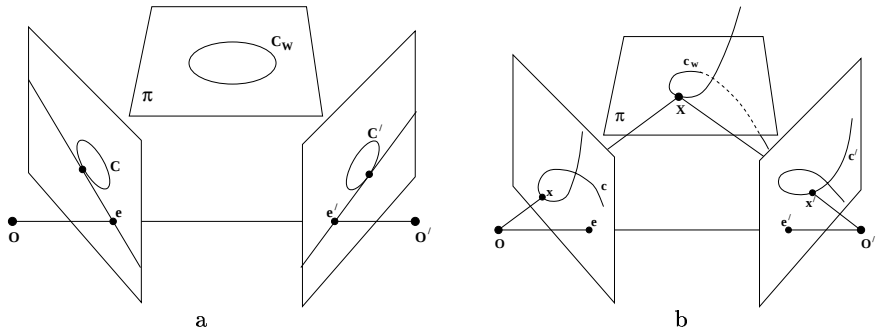


Fig. 1. (a) A conic defines a unique plane in 3-space. This plane is determined up to a two-fold ambiguity from the image of the conic in two views. One of the epipolar tangents is also shown. (b) The osculating plane of a (non-planar) curve varies, but is always defined in 3-space provided the curvature is not zero. This plane is determined uniquely from the image of the curve in two views. In both cases the plane induces a homography between the images.

2.1 Conic planes and homography

In this section we consider the reconstruction of a conic in 3-space given its image in two views and the epipolar geometry of the views represented by the fundamental matrix F. An explicit and simple expression is given for the plane of the conic and the homography it induces. The geometry is illustrated in figure 1a.

Several previous authors have investigated this geometry [12, 17, 19]. The most elegant is the approach of Quan [17] who shows, by reasoning on the back projected cones from each image conic, that there is a two fold ambiguity in the plane of the space conic. However, Quan's expressions are over complicated because he did not explicitly write the two camera matrices in terms of the epipolar geometry. Instead he derived correspondence conditions for the conics

which are, actually, equivalent to the condition that the epipolar tangents to the conic must correspond (as in the Kruppa equations [13]).

We now give the main result of this section:
Given two perspective images of a conic (i.e. not a quadric), and the fundamental matrix between the views, then the plane of the conic (and consequently the homography induced by this plane) is defined up to a two fold ambiguity.
Suppose the image conics are represented by 3×3 symmetric matrices $\mathtt{C}$ and $\mathtt{C}'$ and the 3×4 camera matrices are chosen as

$$\mathtt{P} = [\mathtt{I}|\mathbf{0}] \qquad \mathtt{P}' = [\mathtt{A}|\mathbf{e}']$$

where $\mathbf{e}'$ is the epipole in the second view ($\mathbf{e}$ is the epipole in the first view). Then for the plane $\boldsymbol{\pi}^\top = (\mathbf{a}^\top, 1)$ and homography $\mathbf{x}' = \mathtt{H}\mathbf{x}$ where [11]

$$\mathtt{H} = \mathtt{A} + \mathbf{e}'\mathbf{a}^\top$$

and provided ${\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}' \neq 0$ and $\mathbf{e}^\top \mathtt{C}\mathbf{e} \neq 0$,

$$\mathbf{a}(\mu) = -\left(\mu \mathtt{C}\mathbf{e} + \mathtt{A}^\top \mathtt{C}'\mathbf{e}'\right) / \left({\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}'\right)$$
$$\mathtt{H}(\mu) = \left(\mathtt{I} - \frac{\mathbf{e}'{\mathbf{e}'}^\top \mathtt{C}'}{{\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}'}\right)\mathtt{A} - \mu \frac{\mathbf{e}'\mathbf{e}^\top \mathtt{C}}{{\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}'}$$

and the two values of μ are obtained from

$$\mu^2 \left[(\mathbf{e}^\top \mathtt{C}\mathbf{e})\mathtt{C} - (\mathtt{C}\mathbf{e})(\mathtt{C}\mathbf{e})^\top\right] = \mathtt{A}^\top \left[({\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}')\mathtt{C}' - (\mathtt{C}'\mathbf{e}')(\mathtt{C}'\mathbf{e}')^\top\right] \mathtt{A}$$

If $\mathtt{A}$ is chosen as $\mathtt{A} = [\mathtt{C}'\mathbf{e}']_\times \mathtt{F}$ then simpler expressions are obtained:

$$\mathtt{H}(\mu) = [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} - \mu \mathbf{e}'(\mathtt{C}\mathbf{e})^\top$$

with

$$\mu^2 \left[(\mathbf{e}^\top \mathtt{C}\mathbf{e})\mathtt{C} - (\mathtt{C}\mathbf{e})(\mathtt{C}\mathbf{e})^\top\right] ({\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}') = -\mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F}$$

where the notation $[\mathbf{x}]_\times$ means the 3×3 skew matrix with null space $\mathbf{x}$, representing the vector cross product: $[\mathbf{x}]_\times \mathbf{y} = \mathbf{x} \times \mathbf{y}$.

The proof is omitted for lack of space but simply involves solving for $\mathbf{a}$ in $\mathtt{H} = \mathtt{A} + \mathbf{e}'\mathbf{a}^\top$ such that the image conics map as $\mathtt{C}' = \mathtt{H}^{-\top} \mathtt{C}\, \mathtt{H}^{-1}$. Note that when ${\mathbf{e}'}^\top \mathtt{C}'\mathbf{e}' = \mathbf{e}^\top \mathtt{C}\mathbf{e} = 0$ the epipole is on the conic in each image, and the baseline intersects the space conic. In this case the plane is determined uniquely.

2.2 Curve osculating planes

In this section we first show how Euclidean curvature is mapped by a homography, and then how this result is used to determine the osculating plane of the curve. The geometry is illustrated in figure 1b.

Suppose a plane curve is imaged in two views, then the image curves are related by a homography induced by the curve plane. If the curve is not planar (it

has non-zero torsion) then the image curves are not related globally by a homography, but *locally* they are related by a homography induced by the osculating plane of the space curve.

If corresponding curve points are related as $\mathbf{x}' = \mathtt{H}\mathbf{x}$. Then the Euclidean curvature at these points is related by

$$\kappa' = \kappa \left(\frac{ds}{ds'}\right)^3 \left(\frac{1}{\mathbf{h}_3.\mathbf{x}}\right)^3 |\mathtt{H}| \tag{1}$$

where $|\mathtt{H}|$ is the determinant of $\mathtt{H}$, and $\mathbf{h}_i{}^\top$ is the ith row of $\mathtt{H}$, i.e.

$$\mathtt{H} = \begin{bmatrix} \mathbf{h}_1{}^\top \\ \mathbf{h}_2{}^\top \\ \mathbf{h}_3{}^\top \end{bmatrix}$$

and $\frac{ds}{ds'}$ is the ratio of arc lengths given by

$$\frac{ds'}{ds} = \sqrt{\frac{\left[(\mathbf{h}_1.\dot{\mathbf{x}})\ (\mathbf{h}_3.\mathbf{x}) - (\mathbf{h}_1.\mathbf{x})\ (\mathbf{h}_3.\dot{\mathbf{x}})\right]^2 + \left[(\mathbf{h}_2.\dot{\mathbf{x}})\ (\mathbf{h}_3.\mathbf{x}) - (\mathbf{h}_2.\mathbf{x})\ (\mathbf{h}_3.\dot{\mathbf{x}})\right]^2}{(\mathbf{h}_3.\mathbf{x})^4(\dot{x}_1^2 + \dot{x}_2^2)}}$$

the notation $\dot{\mathbf{x}}$ indicates the derivative of the homogeneous vector $\mathbf{x} = (x_1, x_2, x_3)^\top$ by a curve parameter which need not be arc length and will *not*, in general, be the same in each image. The proof is omitted through lack of space, but starts from the formula of Brill *et al* [14], page 205.

We now show that the known curvatures at corresponding points, $\mathbf{x} \leftrightarrow \mathbf{x}'$, of the curves in each image determine the homography $\mathtt{H}$. Suppose the tangent lines at $\mathbf{x}$, $\mathbf{x}'$, are $\mathbf{l}$ and $\mathbf{l}'$ respectively. These lines are related by $\mathbf{l} = \mathtt{H}^{-\top}\mathbf{l}'$. Given the epipolar geometry, there is a one parameter family of homographies which map a line in one image to a line in the other. This family may be written [20]

$$\mathtt{H}(\mu) = [\mathbf{l}']_\times \mathtt{F} + \mu \mathbf{e}' \mathbf{l}^\top$$

The remaining degree of freedom, μ, is determined *uniquely* using the relation (1) between image curvatures κ' and κ. We introduce the notation $\mathtt{A} = [\mathbf{l}']_\times \mathtt{F}$. Note that since $\mathbf{l} = \mathbf{x} \times \dot{\mathbf{x}}$, it follows that $\mathtt{H}(\mu)\mathbf{x} = \mathtt{A}\mathbf{x}$ and $\mathtt{H}(\mu)\dot{\mathbf{x}} = \mathtt{A}\dot{\mathbf{x}}$, so that $\mathbf{h}_i.\mathbf{x} = \mathbf{a}_i.\mathbf{x}$ etc. After some algebra:

$$\mu\left\{l_1|\mathbf{e}'\mathbf{a}^2\mathbf{a}^3| + l_2|\mathbf{a}^1\mathbf{e}'\mathbf{a}^3| + l_3|\mathbf{a}^1\mathbf{a}^2\mathbf{e}'|\right\} = \frac{\kappa'}{\kappa}\left(\frac{ds'}{ds}\right)^3 (\mathbf{a}_3.\mathbf{x})^3$$

where $\mathbf{a}^i$ is the ith column of $\mathtt{A}$, and $|\mathbf{xyz}|$ is the determinant of the matrix with columns $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$.

To summarise, we have shown that

Given two perspective images of a plane curve, and the fundamental matrix between the views, then the plane of the curve (and consequently the homography induced by this plane) is defined uniquely from the corresponding tangent lines and curvatures at one point.

There are two types of degenerate points at which the plane/homography cannot be uniquely determined by this method. The first is at epipolar tangents, where the tangent lines contain the epipole (i.e. if $\mathbf{l}.\mathbf{e} = \mathbf{l}'.\mathbf{e}' = 0$), since in this case the line in 3-space which projects to $\mathbf{l}$ and $\mathbf{l}'$ can not be determined as it lies in an epipolar plane. The second is at zero's of curvature (inflections) since then the curvature cannot be used to determine μ.

2.3 Transfer to a third view

The previous two sections have shown how the curve plane can be computed from two views. In this section we show that this plane, together with the trifocal tensor of the triplet, defines the mapping of the curve into a third view. This will be illustrated for a conic. The transfer of curvature via the osculating plane follows by analogy. In detail we show that
Given two perspective images of a conic and the trifocal tensor, the image of the conic in the third view is computed up to a two fold ambiguity

Suppose the (known) image conics are $\mathtt{C}$ and $\mathtt{C}'$ in views one and two, and the sought conic in view three is $\mathtt{C}''$. The following steps determine $\mathtt{C}''$

1. Obtain the three $\mathtt{P}$ matrices from the trifocal tensor as described by Hartley [9] $\mathtt{P} = [\mathtt{I}|\mathbf{0}]$, $\mathtt{P}' = [\mathtt{A}|\mathbf{e}']$, $\mathtt{P}'' = [\mathtt{B}|\mathbf{e}'']$.
2. Compute the plane of the homography $\mathbf{a}$ (two solutions) from the corresponding conics $\mathtt{C} \leftrightarrow \mathtt{C}'$ as described in section 2.1.
3. Then the homography $\mathbf{x}'' = \mathtt{H}_{13}\mathbf{x}$ is $\mathtt{H}_{13} = \mathtt{B} + \mathbf{e}''\mathbf{a}^\top$, and the conic $\mathtt{C}$ is transferred as $\mathtt{C}'' = \mathtt{H}_{13}^{-\top}\mathtt{C}\mathtt{H}_{13}^{-1}$.

3 Geometry Assessment

In this section we assess the performance of the methods of section 2 for an implementation on real images. The accuracy is evaluated from three views of a curve as follows: the plane of the curve is estimated from views one and two, and the curve and curvature are then transferred by the induced homography into the third view. The *transfer error* is assessed by comparing the transferred curve/curvature with that of the (veridical) curve imaged in the third view.

Assessment Criteria The transfer is assessed by two error measures. The first, ϵ_{x}, is the average Euclidean distance between the curves

$$\epsilon_{\mathrm{x}} = < d_\perp(\mathbf{c}(s), \mathbf{c}_{\mathrm{t}}(s)) >$$

where $d_\perp$ is the distance between the curves in the normal direction from the imaged curve $\mathbf{c}$ to the transferred curve $\mathbf{c}_{\mathrm{t}}$, and the average is performed over an arc length curve parameter s for the curve $\mathbf{c}$. The correspondence between the

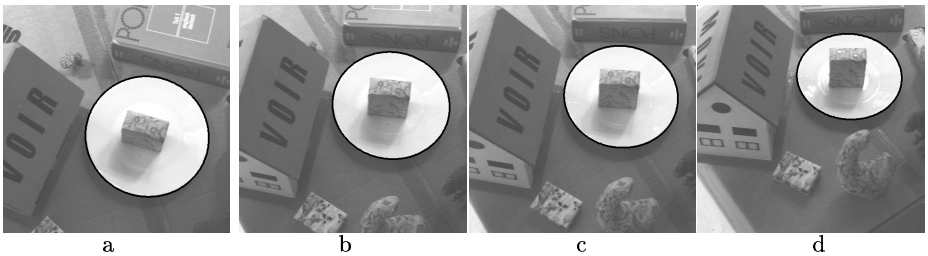
curves is established by intersecting the curve $\mathbf{c}_t$ with the normal of $\mathbf{c}(s)$. The second measure, ϵ_κ, is the average relative curvature difference of the curves

$$\epsilon_\kappa = < \frac{|\kappa(s) - \kappa_t(s)|}{\kappa(s)} >$$

where κ is the curvature of the imaged curve, and κ_t the curvature of the transferred curve. Points at which transfer is not possible, due to degeneracies in determining the osculating plane (see section 2.2), are excluded from the average.

The method is assessed for both conics (i.e. plane curves), and space curves. A conic allows the conic-specific method of section 2.1 to be compared to the general curve method of section 2.2 based on curvature. A conic also allows the curvature method to be assessed in two ways. In the first case a conic is fitted to the extracted conic boundary (edge chain) and intersections, tangents, and curvature are determined from the algebraic curve. This is a *global* method of estimation. In the second case the curve, its tangent, and curvature are determined locally (by spline fitting) using only this *local* information. In general we would expect the global estimates to be superior to the local.

For the plane curves the evaluation images consist of the two image triplets shown in figure 2, with the plate providing a conic. The triplets share the same first two views, but differ in the third. In triplet II the baseline between the second and third views is four times larger than that in triplet I. Space curves are evaluated using the image triplet shown in figure 3. The space curves used for the evaluation are superimposed; one is on the bottle, the second on the mug.



a b c d

Fig. 2. Frames 3, 6, 7 and 10 from the "plate" sequence. The extracted conics are superimposed in black. Test triplet I consists of images (a), (b), (c); and test triplet II of images (a), (b), (d).

Implementation details Conic outlines are extracted using a subpixel Canny edge detector, and the conic estimated by Bookstein's fitting algorithm [3]. The local estimate of curvature is obtained by fitting cubic B-splines to the edgel chains. The average involved in both error measures, ϵ_x and ϵ_κ, is computed by sampling the conic/B-spline curve at one pixel arc length intervals.

Fig. 3. The "bottle" sequence (images 11,15 and 19). Two of the extracted contour chains are superimposed.

3.1 Assessing conic transfer

The conic plane is computed using conics from views one and two and the method of section 2.1. In order to have conics consistent with the epipolar geometry, a correction is required to ensure that the epipolar tangents to the conic (the tangent lines from the epipole to the conic) correspond between the images. This correction is achieved by applying the algorithm of [6]. The corrected conic is then transferred into the third image as described in section 2.3. The distance error and the relative curvature error are computed between the transferred and the extracted conic in the third image. The second solution for the conic plane can easily be ruled out as the distance error is very large.

For triplet I, $\epsilon_{\mathrm{x}} = 0.15$ pixels, and $\epsilon_{\kappa} = 0.0024$. For triplet II, $\epsilon_{\mathrm{x}} = 0.28$ pixels, and $\epsilon_{\kappa} = 0.0074$. Clearly, the transfer is excellent (visually the conics are indistinguishable.).

3.2 Assessing Curvature Transfer

Global curvature estimation Tangent lines and curvatures are obtained for the conics in each view by implicit differentiation of the fitted conics. The correspondence between points on the conic in the first and second view is obtained by intersecting the conic in the second view with the epipolar line of the point in the first view. The conic epipolar tangents and their neighbours are excluded from the error measures since curvature cannot be transferred for these points.

For triplet I, $\epsilon_{\mathrm{x}} = 0.15$ pixels, and $\epsilon_{\kappa} = 0.0016$. For triplet II, $\epsilon_{\mathrm{x}} = 0.34$ pixels, and $\epsilon_{\kappa} = 0.0047$. These results are almost identical to those of the previous section where conic transfer is used (section 2.1), as opposed to conic curvature transfer. They demonstrate that given almost perfect curvature (from the fitted conic) the curvature transfer is similarly almost perfect. The method of the next section investigates the deterioration when the curvature measurement is local and inferior.

Local curvature estimation The conics in image one and two are here represented by B-splines fitted to the edgel chain. The tangents, curvatures and

epipolar line intersections are then computed from the B-splines, as opposed to the conics. The transferred curvature is compared to the curvature computed implicitly from C''. For triplet I, $\epsilon_{\mathrm{x}} = 0.24$ pixels, and $\epsilon_{\kappa} = 0.046$. For triplet II, $\epsilon_{\mathrm{x}} = 0.29$ pixels, and $\epsilon_{\kappa} = 0.074$.

3.3 Comparison of the different methods

The position error ϵ_{x} is of similar value over all methods. This is not surprising because the point transfer error only arises because of the difference between the fitted curves (conic or B-spline) and the veridical conic. This will be consistently sub-pixel because of the sub-pixel acuity of the Canny edge detector used. The curvature error ϵ_{κ} is more illuminating. The error is an order of magnitude larger when using local estimation of curvatures (via B-splines). This increased error is a direct consequence of the error in curvature estimation as can be seen from figure 4 which compares (a) the curvatures measured by the two methods in view one, and then (b) the resulting transferred curvatures in view three.

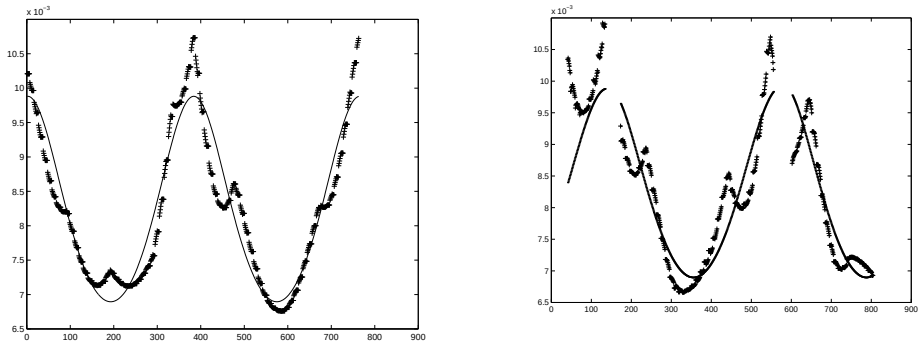


Fig. 4. (a) Comparison of curvature computed from a local B-spline fit (dashed curve), and a global conic fit (continuous curve) to the conic edgel chain. (b) Comparison of the transferred curvature using these two fitting methods for triplet I. In both cases curvature is plotted against arc length in pixels. The missing curve segments are at epipolar tangents where the curvature cannot be transferred.

3.4 Space curves

Figure 3 shows the two space curves used for assessing point and curvature transfer. For the curve on the bottle, $\epsilon_{\mathrm{x}} = 0.43$ pixels, and $\epsilon_{\kappa} = 0.037$. For the curve on the mug, $\epsilon_{\mathrm{x}} = 0.71$ pixels, and $\epsilon_{\kappa} = 0.097$. This is a very similar performance to the case of B-spline curvature computation for plane curves. Again the transfer accuracy is limited by the accuracy of the curvature measurement.

3.5 Rejection of the incorrect solution

Two solutions for the conic plane are obtained in practice using either of the two view methods. In the algebraic method of section 2.1, the two solutions arise because algebraic properties alone do not distinguish the front and back faces of the conic plane [17]. In the local curvature method of section 2.2 the two solutions arise because a curve point in the first view has two potential matches in the second view (since the epipolar line, corresponding to a point on the conic in the first view, intersects the conic in the second view twice). Thus the fallacious solutions originate for different reasons in the two methods. Consequently, the fallacious algebraic solution can be eliminated by simply testing one point with the curvature method: the correct solutions from both methods will be the same, but in general the fallacious solutions differ.

4 Matching curves over multiple views

The algorithm for automatic curve matching uses both photometric information and multiple view constraints. It is an extension to curves of the short and wide baseline algorithms for lines described in [20]. Short and wide base line cases are distinguished by the severity of the differing perspective distortions in the images. In both cases a photometric similarity measure between putatively matched contours is obtained from an aggregated neighbourhood cross-correlation score. In the short baseline case simple neighbourhood cross-correlation of corresponding points suffices; in the wide baseline case corresponding points for the cross-correlation are determined by a local planar projective transformation. The projective transformation is obtained here using the geometric relations of section 2. In the three view case curve matches are also verified using the geometric constraints provided by the trifocal geometry.

4.1 Short baseline between two views

Implementation details Each contour chain is treated as a list of edgels to which neighbourhood correlation is applied as a measure of similarity. The similarity score for a pair of contour chains $\mathbf{c}$ and $\mathbf{c}'$ is computed as the average of the individual edgel correlation values. Point correspondences are determined by the epipolar geometry. In detail, for an image point $\mathbf{x}$ which is an edgel of the contour chain $\mathbf{c}$, the epipolar line in the second image is $\mathbf{l}'_e = \mathtt{F}\mathbf{x}$. This line $\mathbf{l}'_e$ is intersected with the potentially matching edgel chain $\mathbf{c}'$ (using linear interpolation between edgels). The correlation is between the neighbourhood of $\mathbf{x}$ and the neighbourhood of the intersection. Here a correlation window of 15x15 is used. In the case of multiple intersections the best correlation value is included in the overall score, and if there is no intersection, because the chains have different lengths, then no value is added to the score. To be robust to occlusion an individual correlation value is only included if above a threshold (here 0.6). If there are fewer than a minimum number (here 15) of matched edgels for a

putative curve match, then that match is eliminated. No disparity threshold is used. Having determined similarity scores between contour pairs, the matches are selected using a winner take all scheme.

For a given pair of contours which have been deemed matched the parts of the curve for which there are corresponding edgels in both views are determined by only including edgels with correlation above threshold. When moving along a curve figural continuity is used to bridge gaps in the edgel correspondence arising from epipolar tangencies. If the parts of the curve for which there are corresponding edgels are non-contiguous, then only the three longest parts are retained.

Matching performance In the following we present a number of results using the short baseline algorithm. At present the ground-truth matches are assessed by hand. The performance varies with the parameters used for contour extraction and these are now described.

The algorithm used for contour extraction is the 'CannyOx' edge detector and linker from the Targetjr[1] software package. The low and high Canny thresholds for the edge detection are the same in all cases (2 and 12 respectively). The edgel linker parameters are varied, with their values given in table 1. The two parameters are the minimum gradient at which edgels are included in the linked chain (min gradient), — a high value excludes weak edges; and the minimum number of edgels in the linked chain (min length) — a high value excludes short chains. An example of extracted contours using a minimum gradient of 60 and a minimum length of 60 is displayed in figure 5 for the first two bottle images of figure 3.

The matched contours are shown in figure 5 in two ways. First, the complete contour chain matches are shown; second, only the parts of the matched contours for which there are corresponding edgels in both views are shown. The latter choice excludes the parts of the chains along epipolar lines, and also those parts of the chain which are detected as edgels in one view but not in the other. Only corresponding parts are shown for the rest of the examples in this paper.

The number of matched contour chains for different parameters for contour extraction are given in table 1. The matching results are extremely good, with most of the mismatches arising from specularities on the bottle: For the 60/30 case there is one false match, the other three are due to specularities; and for the 30/30 case there are two false matches, the other four are due to specularities. Curves arising from specularities can be removed by a pre-process. It is evident that there is little loss in matching performance when the number of contours is increased, i.e. when the parameter values are decreased.

As a second example contours are matched for the first two images of the "plate" sequence (cf. figure 2). Results are displayed in figure 6. 80 and 89 contours chain are obtained for the left and right image, respectively (using min grad 30 and min length 30); 100% of the 40 matched contour chains are correct.

[1] http://www.esat.kuleuven.ac.be/~targetjr

min grad	min length	number left	number right	number matched	correct
60	60	37	47	29	97%
60	30	59	72	41	90%
30	30	85	85	41	85%

Table 1. Edge detection parameters and curve matching results for the short baseline algorithm applied to an image pair of the "bottle" sequence.

Fig. 5. Short baseline matching. Upper pair : Contours extracted with CannyOx (parameters: min grad 60, min length 60). 37 and 47 contours are obtained for the left and right images respectively; Middle pair : Contours which are matched between views. Lower pair : Matched contours showing only the parts which have corresponding edgels in both views. 97% of the 29 matches are correct.

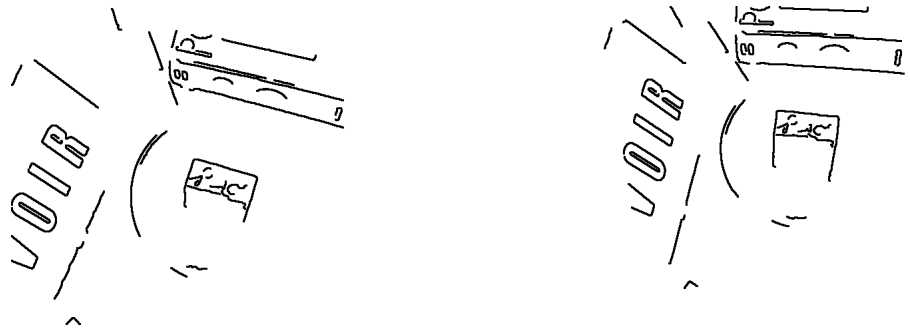


Fig. 6. Short baseline matching. 100% of the 40 matched contour chains are correct.

4.2 Large baseline algorithm between two views

In the case of a large baseline or significant rotation simple correlation based on square neighbourhoods will fail. However, the mapping between neighbourhood points can be determined from a homography based on the local tangent plane of the surface. In the case of line matching [20] this homography can only be determined up to a one parameter family because a line in 3D only determines the plane inducing the homography up to a one parameter family. This means that for lines a one dimensional search is required. However, in the case of curves the osculating plane determines a local plane uniquely (or up to a two-fold ambiguity for conics), and no search is required.

Thus there are two cases: if the curve has sufficient curvature the plane is computed as described in section 2.2; however, if the contour chain is straight or almost straight, curvature is close to zero and the osculating plane cannot be determined. In this case the algorithm developed for long range motion line matching (cf. [20]) is used. Lines here are the tangent lines to the contour chain. A threshold on local curvature is used to decide which case applies.

Images 11 and 19 of the "bottle" sequence (cf. figure 3) are matched; results are displayed in figure 7. 37 and 48 contour chains were extracted for the left and right images, respectively. 88% of the 16 matched contour chains are correct.

Fig. 7. Wide base line matching. 88% of the 16 matched contour chains are correct.

4.3 Three view matching

In the three view case there are three geometric constraints available for matching curves [18]: the curve point, its tangent, and its curvature. Here only the point is used directly as a geometric constraint. The tangent and curvature are used indirectly to determine the osculating plane, and hence provide a strong photometric constraint based on neighbourhood cross-correlation via the homography induced by the plane.

The three view matching builds on the curve matches provided by the two view algorithm. The point correspondences for matched chains in the first and second view are transferred into the third view using the trifocal tensor and are checked for coincidence with a contour chain in the third image (geometric constraint). A photometric score is then computed by correlating the point pair between images 2 and 3. If sufficient corresponding point triples are obtained for a putative contour triple, its mean correlation score is used to rank the match; otherwise the putative match is eliminated. This score is used in a winner take all scheme over three views which rules out multiple matches.

In the following results are given for two image triplets: "bottle" (images 11, 15, 17 of figure 3), and "plate" (images 3, 6, 7 of figure 2). Results on other triplets are given in [21]. Contours are extracted using the same parameters in all cases (min grad 30, min length 30).

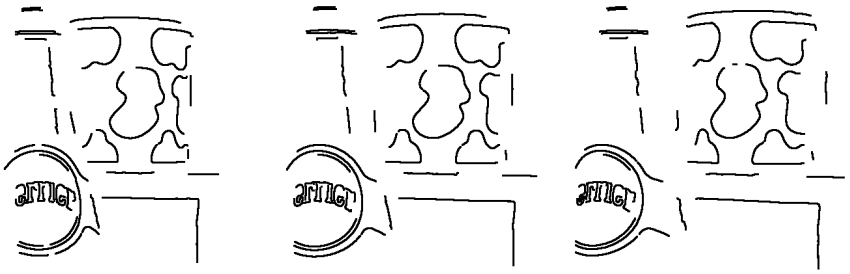


Fig. 8. Three view matching. 25 contour chains are matched. There is only one incorrect match which arises from a specularity on the bottle.

Figure 8 shows results for three view matching for the "bottle" images (cf. figure 3). 85, 85 and 90 contour chains are obtained for the left, middle and right image, respectively. 25 contour chains were matched. Note, that the outline of the bottle is correctly *not* matched over the three views since it is an apparent contour and not rigidly attached to the scene. Annoyingly, the non-rigidity was insufficient to eliminate the only incorrect match which arose from the specularity on the bottle. In figure 9 matched contours are displayed for the "plate" images (cf. figure 2). 80, 89 and 84 contour chains are obtained for the left, middle and right image, respectively. 100% of the 30 matched contour chains are correct.

The three view matching gives excellent results. These matches will now form the foundation for a second pass where curves broken by the contour extraction

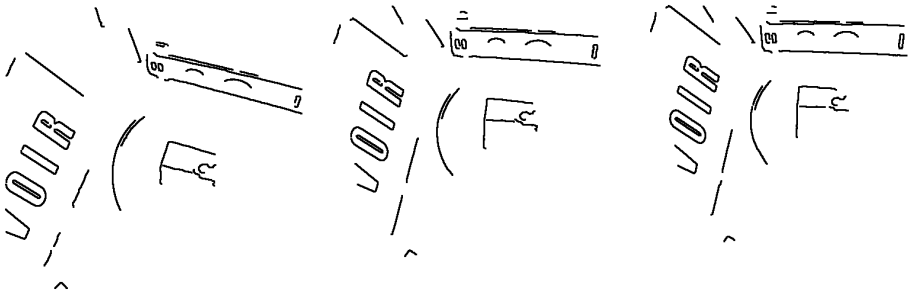


Fig. 9. Three view matching. 100% of the 30 matched contour chains are correct.

can be rejoined, and edge evidence sought for curve portions which are matched in only two of the three views. Such post processing substantially reduces curve fragmentation [20].

5 Discussion and extensions

We have derived and implemented new geometric results for conics and general curves over multiple views. These results have been applied to automatic curve matching over two and three views using a combination of geometric and photometric constraints.

The short baseline algorithm gives substantial disambiguation and generally excellent results being extremely robust with three views. The wide baseline algorithm also gives very good results, but relies on computing the osculating plane of the curve in order to determine a homography to map intensity neighbourhoods for correlation. There are two potential problems here: first, the osculating plane is only estimated accurately if the torsion of the curve is sufficiently small. However, this condition holds in many practical cases; second, the osculating plane of the curve may not coincide with the local tangent plane of the surface (think of the circular curve at the mouth of a mug). Fortunately, the intensity cross correlation is quite forgiving, and provides disambiguation even if the plane is imprecise.

Acknowledgements
We are very grateful to Andrew Fitzgibbon for both discussions and software. Financial support for this work was provided by EU Esprit Project IMPACT.

References

1. P. Beardsley, P. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *Proc. ECCV*, pages 683–695, 1996.
2. A. Blake and R. Cipolla. Robust estimation of surface curvature from deformation of apparent contours. In *Proc. ECCV*, pages 465–474, 1990.
3. F.L. Bookstein. Fitting conic sections to scattered data. In *Computer Graphics and Image Processing*, 9:56–71, 1979.
4. A.T. Brint and M. Brady. Stereo matching of curves by least deformation. In *International Workshop on Intelligent Robots and Systems*, pages 163–170, 1989.

5. R. C. K. Chung and R. Nevatia. Use of monocular groupings and occlusion analysis in a hierarchical stereo system. In *Proc. CVPR*, pages 50–56, 1991.
6. G. Cross and A. Zisserman. Quadric surface reconstruction from dual-space geometry. In *Proc. ICCV*, pages 25–31, 1998.
7. O. Faugeras and L. Robert. What can two images tell us about a third one? *International Journal of Computer Vision*, 18:5–19, 1996.
8. R.I. Hartley. A linear method for reconstruction from lines and points. *Proc. ICCV*, pages 882–887, 1995.
9. R.I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125–140, 1997.
10. P. Havaldar and G. Medioni. Segmented shape-descriptions from 3-view stereo. *Proc. ICCV*, pages 102–108, 1995.
11. Q. Luong and T. Vieville. Canonic representations for the geometries of multiple projective views. Technical report, University of California, Berkeley, 1993.
12. S. Ma. Conics-based stereo, motion estimation, and pose determination. *International Journal of Computer Vision*, 10(1):7–25, 1993.
13. Maybank S. and Faugeras O. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.
14. J.L. Mundy and A. Zisserman, editors. *Geometric Invariance in Computer Vision*. The MIT Press, Cambridge, MA, USA, 1992.
15. S.K. Nayar and R.M. Bolle, Reflectance Based Object Recognition. *International Journal of Computer Vision*, 17(3):219–240, 1996.
16. S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient constraint. *Perception*, 14:449–470, 1985.
17. L. Quan. Conic reconstruction and correspondence from two views. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 18(2):151–160, 1996.
18. L. Robert and O.D. Faugeras. Curve-based stereo: Figural continuity and curvature. In *Proc. CVPR*, pages 57–62, 1991.
19. R. Safaee-Rad, I. Tchoukanov, B. Benhabib, and K.C. Smith. 3D pose estimation from a quadratic curved feature in two perspective views. In *Proc. ICPR*, pages 341–344, 1992.
20. C. Schmid and A. Zisserman. Automatic line matching across views. In *Proc. CVPR*, pages 666–671, 1997.
21. C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. Research Report, 1998.
22. A. Shashua. Trilinearity in visual recognition by alignment. In *Proc. ECCV*, pages 479–484, 1994.
23. M. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *International Journal of Computer Vision*, pages 171–183, 1990.
24. P. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*, 15:591–605, 1997.
25. R. Vaillant. Using occluding contours for 3D object modeling. In *Proc. ECCV*, pages 454–464, 1990.
26. Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78:87–119, 1995.
27. Y. Zhang and J.J. Gerbrands. Method for matching general stereo planar curves. *Image and Vision Computing*, 13(8):645–655, 1995.
28. Zisserman A. and Maybank S. A case against epipolar geometry. In *Applications of Invariance in Computer Vision LNCS 825*. Springer-Verlag, 1994.